



OPEN

## Determinants of vitamin D status in Kenyan calves

Rebecca Callaby<sup>1,2</sup>, Emma Hurst<sup>3</sup>, Ian Handel<sup>1</sup>, Phil Toye<sup>4</sup>, Barend M. de C. Bronsvort<sup>1,2</sup> & Richard J. Mellanby<sup>3</sup>✉

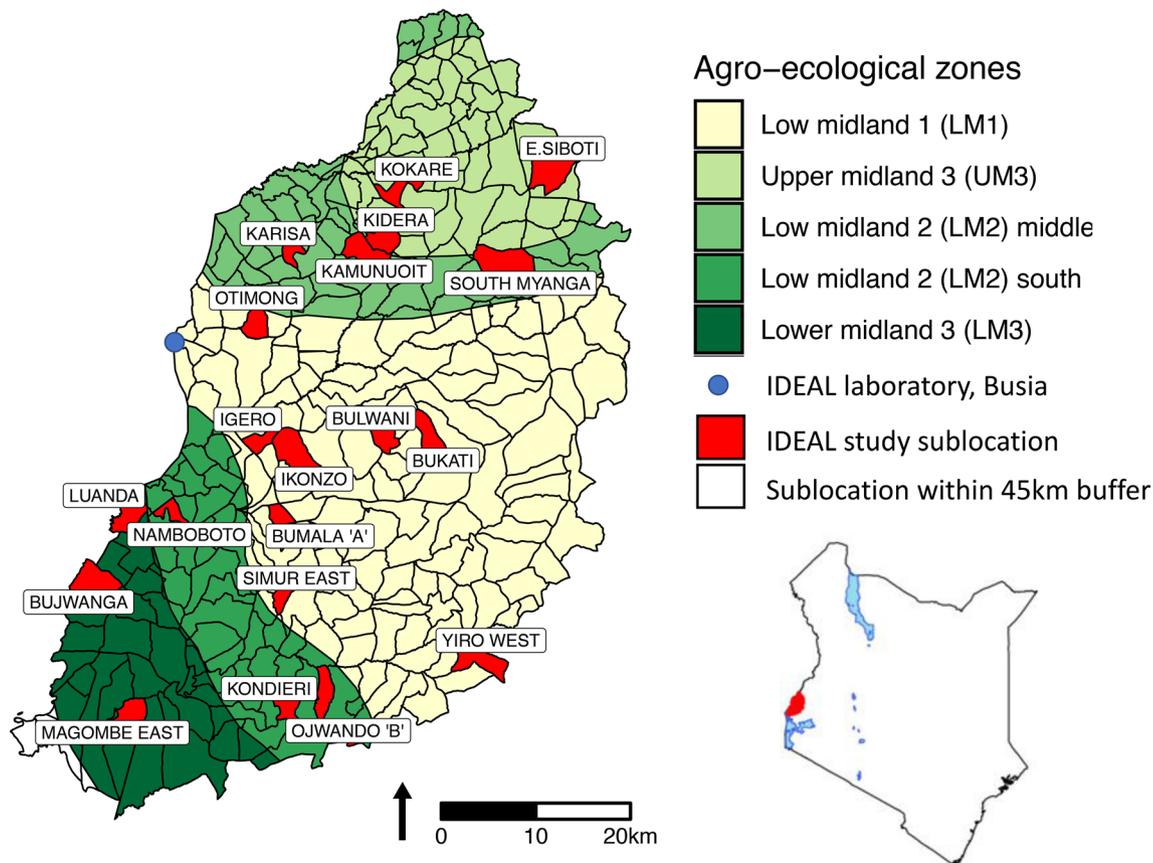
Vitamin D plays a critical role in calcium homeostasis and in the maintenance and development of skeletal health. Vitamin D status has increasingly been linked to non-skeletal health outcomes such as all-cause mortality, infectious diseases and reproductive outcomes in both humans and veterinary species. We have previously demonstrated a relationship between vitamin D status, assessed by the measurement of serum concentrations of the major vitamin D metabolite 25 hydroxyvitamin D (25(OH)D), and a wide range of non-skeletal health outcomes in companion and wild animals. The aims of this study were to define the host and environmental factors associated with vitamin D status in a cohort of 527 calves from Western Kenya which were part of the Infectious Disease of East African Livestock (IDEAL) cohort. A secondary aim was to explore the relationship between serum 25(OH)D concentrations measured in 7-day old calves and subsequent health outcomes over the following 12 months. A genome wide association study demonstrated that both dietary and endogenously produced vitamin D metabolites were under polygenic control in African calves. In addition, we found that neonatal vitamin D status was not predictive of the subsequent development of an infectious disease event or mortality over the 12 month follow up period.

Vitamin D plays a critical role in calcium homeostasis and in the maintenance and development of skeletal health. Vitamin D can be obtained from dietary sources or from production in the skin<sup>1</sup>. Cattle can obtain vitamin D from ingestion of vitamin D<sub>2</sub> or D<sub>3</sub><sup>2</sup>. Vitamin D<sub>2</sub> is present in some plants following the conversion of ergosterol to vitamin D<sub>2</sub> by ultraviolet radiation. Cattle are only likely to consume dietary sources of vitamin D<sub>3</sub> if they have access to proprietary foodstuffs which are directly supplemented with vitamin D<sub>3</sub> or contains vitamin D<sub>3</sub> rich ingredients such as oily fishes. Vitamin D<sub>3</sub> can be produced cutaneously following the isomerisation of 7-dehydrocholesterol by ultraviolet radiation<sup>3</sup>. If insufficient vitamin D is consumed or produced cutaneously, skeletal complications can develop in cattle as a sequelae of prolonged and severe vitamin D deficiency<sup>2</sup>.

Although the importance of vitamin D has been recognised for several decades in the development and maintenance of skeletal health<sup>1,4</sup>, the role of vitamin D in non-skeletal health outcomes has been extensively explored following the discovery that many cell types express the vitamin D receptor<sup>5,6</sup>. Vitamin D metabolites have subsequently been shown to have extensive immunomodulatory roles in experimental models and human studies<sup>7-9</sup>. Low vitamin D status has been linked to all-cause mortality<sup>10-12</sup> and numerous other infectious<sup>13-15</sup>, allergic<sup>16</sup> and autoimmune diseases<sup>17,18</sup> in humans. Vitamin D supplementation trials have indicated potential beneficial roles in cancer outcomes<sup>19</sup> and reduction in diabetes risk in vitamin D deficient individuals<sup>20</sup>. Similarly in companion animals, low vitamin D status has been linked to mortality in hospitalised cats<sup>21</sup> and dogs<sup>22</sup> and in animals with specific diseases such as dogs with chronic enteropathies<sup>23,24</sup>. Low vitamin D status has also been associated with inflammation in dogs<sup>25</sup> and cats<sup>26</sup>.

The relationship between vitamin D and non-skeletal health outcomes has been less studied in ruminants. We have previously reported a positive association between autumnal serum 25(OH)D concentrations and subsequent fecundity<sup>27</sup> and birth weight<sup>28</sup> the following spring in sheep. Vitamin D metabolites have been shown to modulate bovine immune cell phenotype *in-vitro*<sup>29-32</sup> and *in-vivo*<sup>33,34</sup>. In addition, vitamin D status has been found to be significantly lower in cattle with clinical *Mycobacterium avium paratuberculosis* disease compared to cows in a subclinical stage and non-infected control cows further supporting a potential role of vitamin D in regulating non-skeletal health outcomes<sup>35</sup>.

<sup>1</sup>The Epidemiology, Economics and Risk Assessment (EERA) Group, The Roslin Institute and The Royal (Dick) School of Veterinary Studies, Easter Bush Veterinary Centre, The University of Edinburgh, Roslin EH25 9RG, Midlothian, UK. <sup>2</sup>Centre for Tropical Livestock Genetics and Health (CTLGH), The Roslin Institute, University of Edinburgh, Easter Bush Campus, Roslin EH25 9RG, UK. <sup>3</sup>The Vitamin D Animal Laboratory (VitDAL), The Royal (Dick) School of Veterinary Studies and The Roslin Institute, Easter Bush Veterinary Centre, The University of Edinburgh, Roslin EH25 9RG, Midlothian, UK. <sup>4</sup>International Livestock Research Institute and Centre for Tropical Livestock Genetics and Health, Nairobi, Kenya. ✉email: Richard.Mellanby@ed.ac.uk



**Figure 1.** Map of western Kenya showing the study area, agro-ecological zones and sublocations. Selected sublocations are highlighted in red. Figure taken from Bronsvoort et al. (2013).

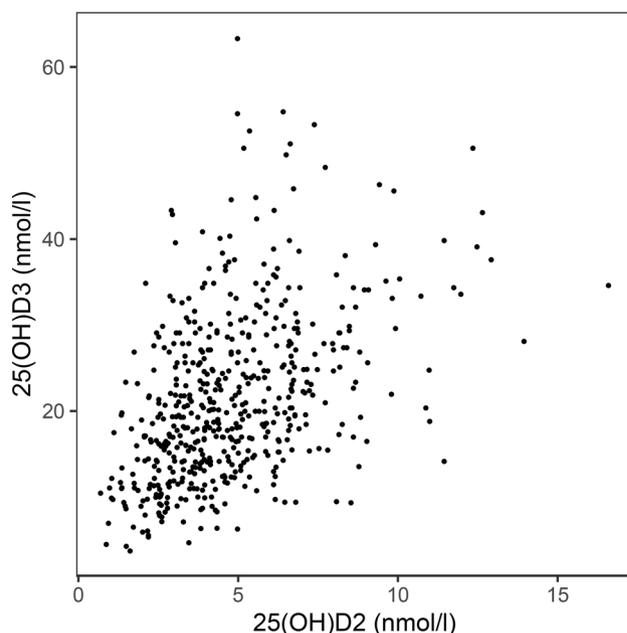
Despite the accumulating evidence that vitamin D may regulate non-skeletal health outcomes in cattle, little is understood about the genetic and environmental factors which are associated with vitamin D status. Furthermore, no studies have explored the relationship between neonatal vitamin D status and subsequent health outcomes in a population of large, healthy calves in a non-experimental, natural setting. We have addressed these research questions by measuring 25(OH)D concentrations in 527 calves from Western Kenya which were part of the precisely phenotyped Infectious Disease East African Livestock (IDEAL) cohort<sup>36</sup>, see Fig. 1 for a map of the study area. All calves in this cohort were examined by an animal health technician every 5 weeks and detailed information about infectious disease incidence was recorded. All calves were followed until 12 months of age or until death. This cohort provided us with an excellent opportunity in which to explore vitamin D regulation in cattle and its relationship with subsequent health outcomes.

## Results

**Cohort description.** In total, serum total 25(OH)D, 25(OH)D<sub>2</sub> and 25(OH)D<sub>3</sub> concentrations were available from 527 calves which were included in this study. The cattle were all between 3–7 days old, East African Shorthorn Zebus and there were 278 males and 249 females. The cattle had a mixture of coat colours, 63 having a light coat, 88 having a dark coat, 174 having a brown coat, 194 having a mixed coat. Eight cattle did not have their coat colour recorded. In addition, 428 of the farmers enrolled in this study stated that they gave cattle in their herd nutritional supplements. Furthermore, the dams of the studied calves were mainly of normal body condition score (n = 366), or fat (n = 145), very few dams were lean (n = 16).

Serum 25(OH)D<sub>2</sub> and 25(OH)D<sub>3</sub> concentrations were positively correlated (Spearman rank correlation = 0.45,  $p < 0.001$ , Fig. 2). On average, 25(OH)D<sub>3</sub> concentrations were 16.71 nmol/l (95% CI = 15.94–17.47 nmol/l) higher than 25(OH)D<sub>2</sub> concentrations (Table 1).

**Association with calf level factors.** Graphs showing serum 25(OH)D concentrations and calf level variables of interest are presented in Figs. 3 and 4 and supplementary Fig. 1. After accounting for AEZ, calves with darker coat colours had lower serum 25(OH)D<sub>3</sub> and total 25(OH)D concentrations than calves with mixed coat colours (Fig. 5); whereas there was no association between calf coat colour and 25(OH)D<sub>2</sub> concentration. Furthermore, calves in upper midland 3 (UM3) had lower serum concentrations of 25(OH)D<sub>3</sub> than calves in low midland 1 (LM1) however there was no association between AEZ and 25(OH)D<sub>2</sub> or total 25(OH)D concentrations (Fig. 5). In addition, nutritional feeding and calf gender was not associated with any of the serum 25(OH)D concentrations (Fig. 5).



**Figure 2.** Scatterplot showing the relationship between 25(OH)D<sub>2</sub> and 25(OH)D<sub>3</sub>.

**Association with environmental level factors.** After accounting for calf coat colour and AEZ, there was no association between serum 25(OH)D concentration and average cloud cover or number of hours of sunshine per day in the month prior to the visit (Fig. 6). There was however an association between serum 25(OH)

	Mean (nmol/l)	Standard error	Range (nmol/l)	Median (nmol/l)	Inter quartile range (nmol/l)
25(OH)D <sub>2</sub>	4.82	0.10	0.72–16.62	4.42	2.95
25(OH)D <sub>3</sub>	21.53	0.42	3.82–63.40	19.79	12.90
Total 25(OH)D	26.36	0.48	5.46–68.42	24.66	13.77

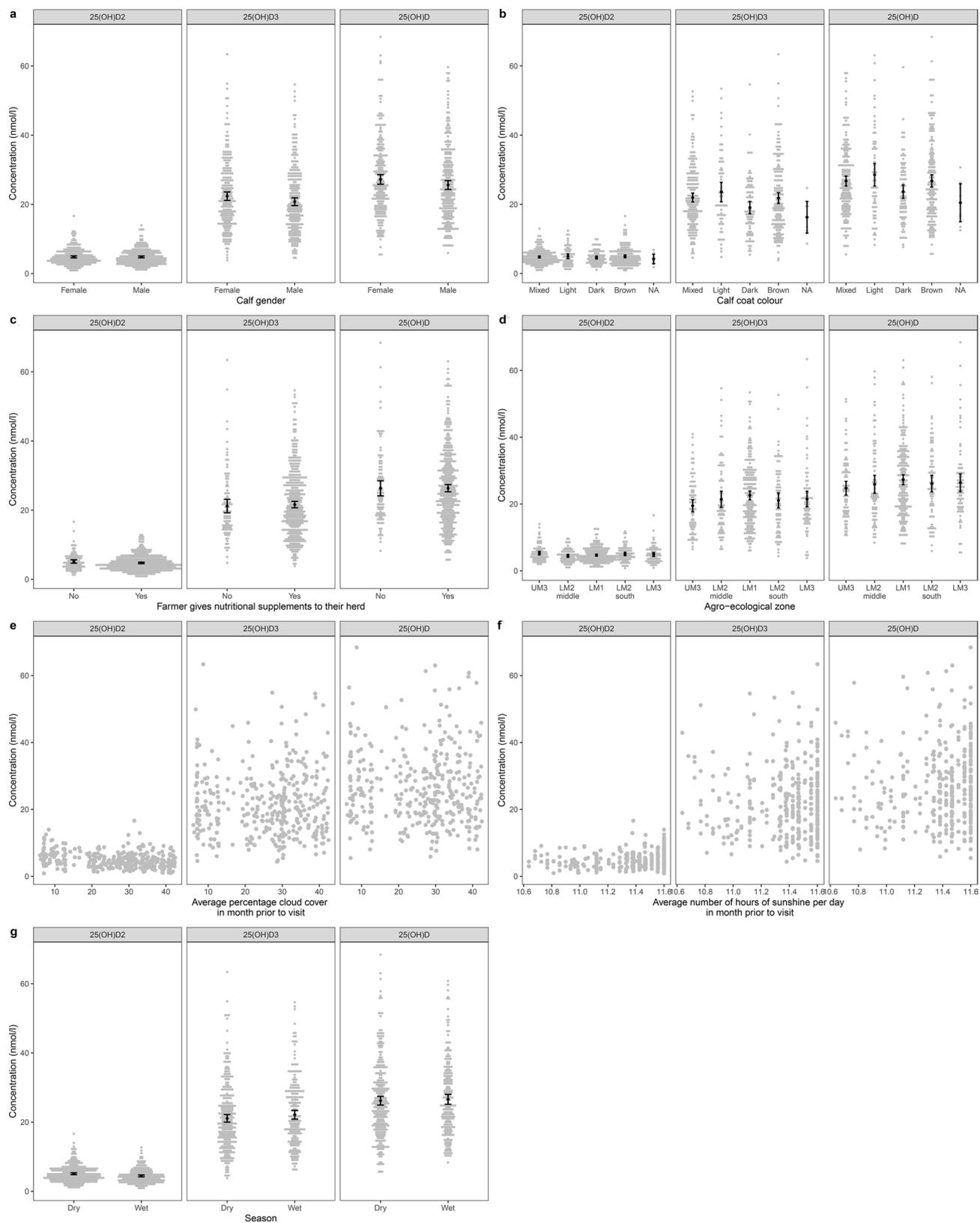
**Table 1.** Descriptive summary of the mean serum 25(OH)D concentrations.

D<sub>2</sub> concentration and season, cattle which were born during the wet season had lower serum 25(OH)D<sub>2</sub> concentrations than those born during the dry season (Fig. 7).

**Association with infectious disease mortality and morbidity.** Out of the 527 cattle enrolled in this study, 81 (15.4%) individuals died as a result of infectious causes before the age of 1 year. A further nine individuals died of non-infectious causes such as trauma, cassava poisoning or mis-mothering and were censored from the study. In addition, 284 (53.89%) calves experienced a clinical episode during the course of the study. The median time to first clinical illness was 106 days (interquartile range = 122 days).

After accounting for calf level factors and AEZ, there was no association between serum 25(OH)D concentrations and mortality from infectious disease or ever experiencing a clinical episode (Figs. 8 and 9). Furthermore, there was no association between serum 25(OH)D concentration and time to first clinical episode (Figs. 8 and 9). In addition, there was no association between serum 25(OH)D concentrations and host outcome when serum 25(OH)D concentrations were included as a categorical value based on tertiles.

**Heritability and genome wide association study of Vitamin D.** The crude heritability of vitamin D status, after accounting for confounding from calf sex, calf coat colour, nutritional supplements and AEZ, is 25(OH)D<sub>2</sub>  $h^2 = 62.6\%$  (SE = 35.3); 25(OH)D<sub>3</sub>  $h^2 = 70.1\%$  (SE = 30.2) and total 25(OH)D  $h^2 = 71.1\%$  (SE = 30.0), respectively (Table 2). One SNP which was associated with 25(OH)D<sub>3</sub> and total 25(OH)D with a p-value greater than the suggestive threshold of  $< 1^{-5}$  (Fig. 10). No SNPs were associated with 25(OH)D<sub>2</sub> at this level. This SNP was located at Chr 3:104814609 (Fig. 10). Ensembl Release 99<sup>37</sup> identified one gene within  $\pm 5000$  bp of this SNP; the gene was FOXO6 which is a member of the FoxO class of transcription factors with distinct shuttling dynamics.



**Figure 3.** Graphs showing serum 25(OH)D concentration and calf and environmental level variables of interest. Black points represent the mean with 95% confidence interval. **(a)** Calf gender; **(b)** calf coat colour; **(c)** nutritional supplements; **(d)** agro-ecological zone; **(e)** average percentage cloud cover in month prior to visit; **(f)** average number of hours of sunshine per day in month prior to visit; **(g)** wet or dry season.

## Discussion

Our study discovered that host factors are associated with vitamin D status in cattle. Similar to our observations in sheep in temperate regions<sup>27,28</sup>, we found that dark coat colour was associated with lower serum 25(OH)D<sub>3</sub> concentrations in African calves which we would predict is a result of lower cutaneous production of vitamin D. Our study failed to identify clear genetic loci associated with serum 25(OH)D concentrations in cattle. This observation mirrors findings in humans which have also shown that total 25(OH)D concentration is under polygenic control with only a modest number of genetic loci linked to vitamin D status<sup>38,39</sup>. A GWAS study of nearly 80,000 individuals identified only six loci linked to 25(OH)D concentrations resulting in an estimate of heritability of serum 25(OH)D concentrations attributable to GWAS common SNPs at only 7.5%<sup>38</sup>. A more recent GWAS study of 25(OH)D concentrations involving over 400,000 participants has identified a further 63 loci but the SNP heritability of 25(OH)D was estimated to be 16.1%<sup>39</sup>. We only found one SNP associated with vitamin D status in this study which may reflect the relatively low coverage of the SNP array and the number of individuals in the analysis.

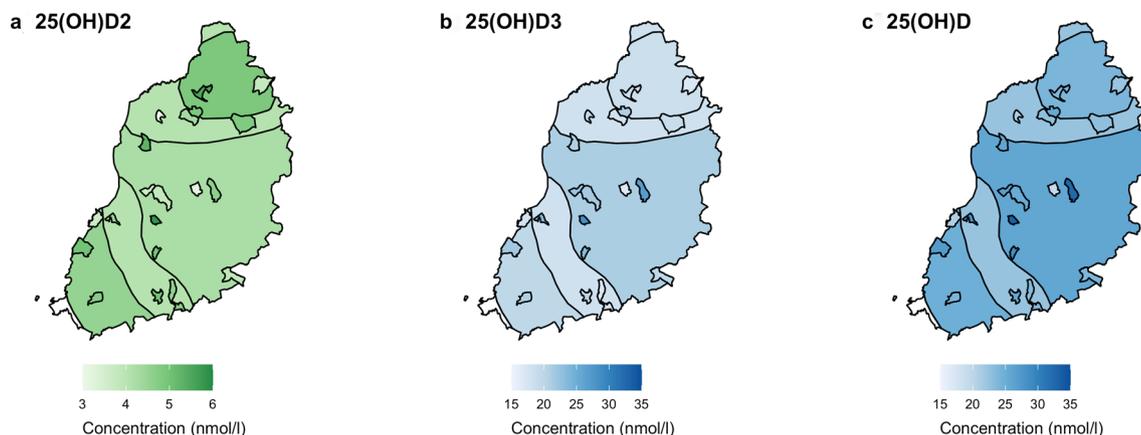
The lack of association between vitamin D status and mortality contrasts with cross sectional studies in humans and companion animals<sup>11,12,21,22,24,40</sup>. It also contrasts with some, but not all, mendelian randomisation studies which have linked increased risk of mortality in humans with genetically lower vitamin D status<sup>10,41–43</sup> and with findings that vitamin D supplementation lowers mortality in patients with cancer<sup>19</sup>. The explanation for the discordant result in our study is unclear. A potential explanation is that vitamin D is simply a marker of ill health and a low vitamin D status is a consequence, rather than a cause, of ill-health. This concept is supported by studies which reveal that 25(OH)D concentrations decline following the development of an inflammatory response<sup>44,45</sup>. Similarly, cattle infected with experimental bovine diarrhoea virus experienced a significant decrease in 25(OH)D concentrations<sup>46</sup>. Another potential explanation is that young animals have lower vitamin D status than adults and consequently the range of 25(OH)D concentrations is smaller than in adults<sup>27,47</sup>. The small variation in 25(OH)D concentrations in the IDEAL calves may reduce the study's ability to detect relationships between 25(OH)D concentrations and longer-term health outcomes. In addition, the higher infectious disease challenge faced by calves in Kenya may limit the impact of vitamin D to shape non-skeletal health<sup>34</sup>. Finally, 25(OH)D concentrations measured in different years tend to show little variation in adult human studies<sup>48–52</sup> but neonatal vitamin D status invariably cannot absolutely predict future 25(OH)D concentrations<sup>53</sup>. Consequently, further studies should explore whether 25(OH)D concentrations from adult African cattle are predictive of mortality.

We found no evidence that vitamin D status was linked to subsequent infectious disease clinical episode events. This contrasts with findings in humans where low cord or neonatal serum 25(OH)D concentrations have been associated with adverse health outcomes such as early childhood lung disorders<sup>54,55</sup>, higher blood pressure<sup>56</sup>, atopic dermatitis<sup>54</sup>, food allergy<sup>53</sup> and sepsis<sup>57,58</sup>. However, our finding is consistent with findings in experimental studies of respiratory infections in cattle. Calves with low vitamin D status did not have any significant differences in lung pathology following challenge with respiratory syncytial virus compared to calves with high vitamin D status<sup>59</sup>. Although vitamin D metabolites can dramatically influence the bovine immune response both *in-vitro*, the impact on the host immune system following *in-vivo* vitamin D supplementation is less consistent. For example, vitamin D repletion had no impact on a wide range of immune parameters in cattle<sup>60</sup> and so the differences in vitamin D status in the cattle included in this study may be insufficient to modulate the immune system in a clinically detectable manner. Another potential explanation is that the calves, who resided in a tropical location, were exposed to sufficient UVB radiation to prevent the development of profound vitamin D insufficiency<sup>61–63</sup>. Numerous, albeit not all, studies in humans have shown that patients with the lowest vitamin D status have the greatest risk of disease development<sup>14</sup> and achieve the most significant health benefits from vitamin D supplementation<sup>64,65</sup>.

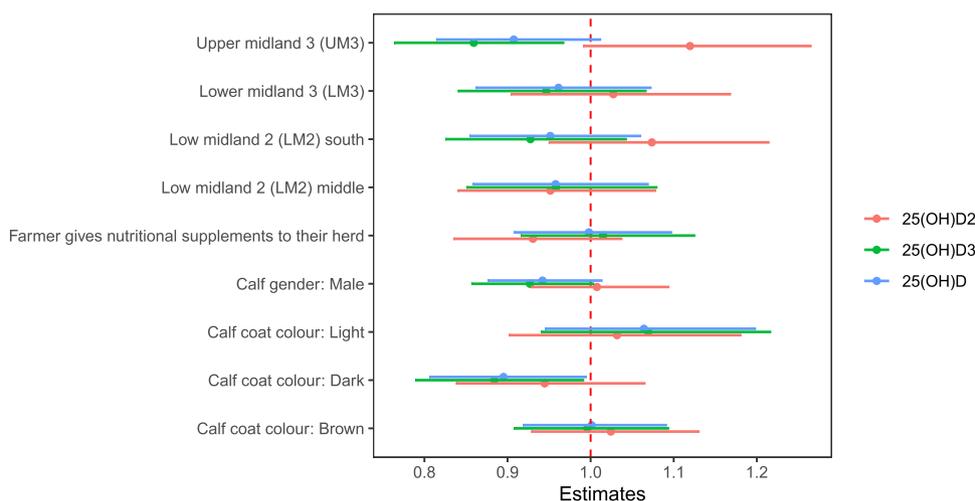
In summary, we have shown that vitamin D status is under polygenic control in African cattle and that host factors, namely coat colour, are associated with serum 25(OH)D<sub>3</sub> concentrations. We found no evidence that neonatal vitamin D status is associated with infectious diseases development or mortality in African calves. This work further illuminates the ongoing debate about the non-skeletal health benefits of vitamin D in production animals. Further studies exploring the relationship between neonatal vitamin D status and future health outcomes in non-tropical populations will be informative. In addition, studies exploring the relationship between 25(OH)D concentrations in young adult African cattle and infectious disease development and overall mortality will also further inform the discussion around the non-skeletal health benefits of vitamin D in ruminants.

## Methods

**Study population.** The Infectious Disease East African Livestock (IDEAL) project was a longitudinal cohort study which was conducted in Western Kenya between 2007–2009; see Bronsvort et al. (2013) for a detailed description of the study design<sup>36</sup>. Briefly, 548 indigenous East African Shorthorn Zebu calves in were randomly selected using a stratified two-stage clustered study design. In the first stage, 20 sublocations (the smallest administrative unit in Kenya) were selected from five agro-ecological zones (AEZ), across an area of roughly 45 × 90 km<sup>2</sup> surrounding the town of Busia, Fig. 1. In the second stage, around 28 calves were recruited from each sublocation. Only one calf was recruited per farm. Calves were between 3 and 7 days old at recruitment and were followed for the first year of life. They were visited every 5 weeks for a clinical examination by a team of veterinarians and animal health assistants. If the farmer reported that the calf was experiencing a clinical episode, then an interim visit was carried out. If the calf died, then a post-mortem examination was performed to attribute the cause of death. Blood and tissue samples were collected in association with all visits and screened using a range of laboratory based diagnostic methods for over 100 different pathogens or infectious exposures. Information of the dams of the calves was also collected at each 5-weekly visit until the calf was weaned. This included information on the body condition score of the dam, which was measured on a scale of 1–9 with 1–3



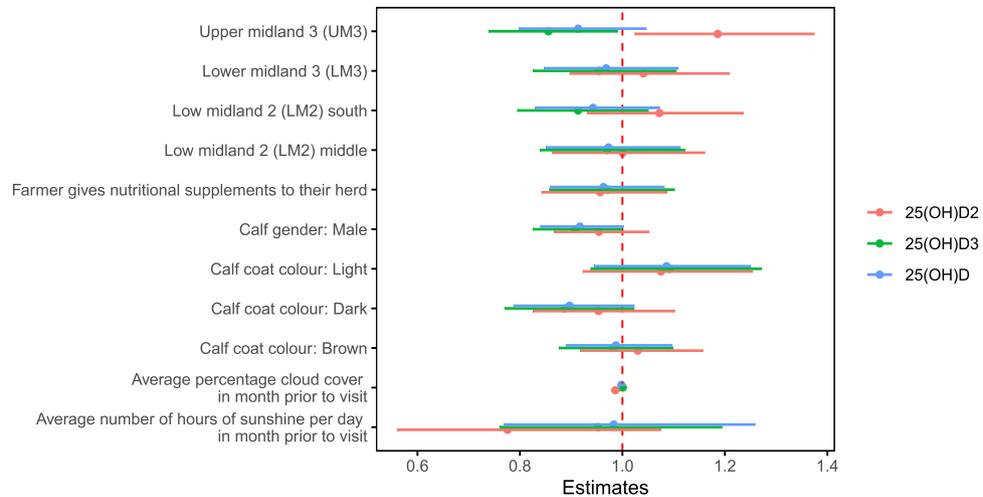
**Figure 4.** Median serum 25(OH)D concentration (nmol/l) in each agro-ecological zone and sublocation. Figures a–c represent: (a) 25(OH)D<sub>2</sub>; (b) 25(OH)D<sub>3</sub> and (c) 25(OH)D. Refer to Fig. 4 for the names of the agro-ecological zones and sublocations.



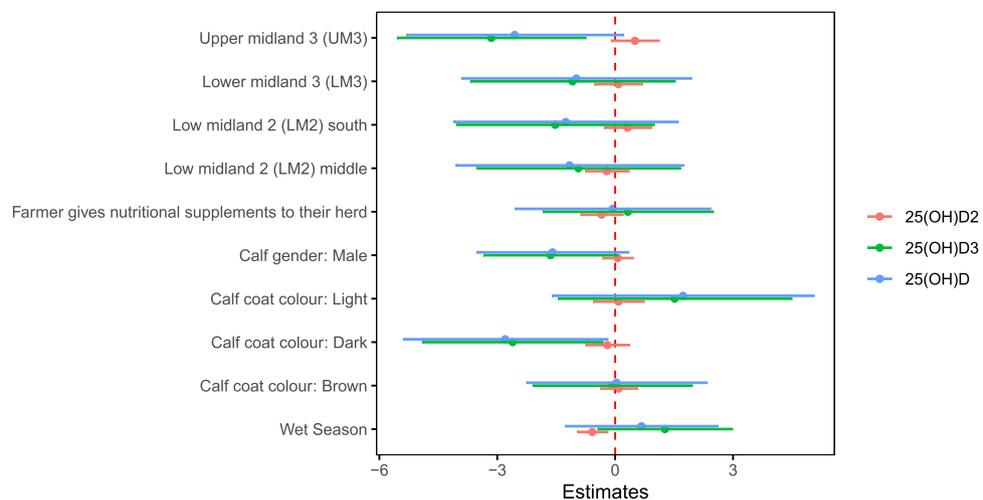
**Figure 5.** Estimates of effect sizes for the calf level factors affecting serum 25(OH)D concentration. Average model estimates are calculated for variables in the top model set (models with a cumulative Akaike weight  $\leq 0.95$ , Supplementary Table 1 and Supplementary Fig. 2). Error bars represent the 95% confidence intervals. Factors are considered important if their confidence intervals do not span one, as indicated by the dotted vertical line.

meaning lean, 4–6 meaning normal weight and 7–9 being fat<sup>66</sup>. All the data collected during the IDEAL is stored in a comprehensive open-access database, in addition, the samples collected during the IDEAL project are stored in a bio-repository at  $-80\text{ }^{\circ}\text{C}$ <sup>67</sup>.

**Measuring serum 25(OH)D concentration.** In 2018, the serum concentration of 25(OH)D<sub>2</sub> and 25(OH)D<sub>3</sub> from bio-banked samples taken from the 548 IDEAL calves at the recruitment visit (when the cattle were 3–7 days old) was measured as previously described<sup>68</sup>. Serum concentrations of 25(OH)D<sub>2</sub> and 25(OH)D<sub>3</sub> were measured by liquid chromatography tandem mass spectrometry (LC–MS/MS) by the Vitamin D Animal Laboratory (VitDAL) which has been certified as proficient by the international Vitamin D Quality Assurance Scheme (DEQAS). Briefly, 200  $\mu\text{L}$  of serum and 1% BSA calibrators spiked with certified reference standards 25(OH)D<sub>2</sub> and 25(OH)D<sub>3</sub> (Sigma-Aldrich, UK) were spiked with equal amounts of isotopically labelled internal standards (deuterium-labeled 25(OH)D<sub>2</sub> (6,19,19-d<sub>3</sub>-25(OH)D<sub>2</sub>) and carbon-13-labeled 25(OH)D<sub>3</sub> (23,24,25,26,27-<sup>13</sup>C<sub>5</sub>-25(OH)D<sub>3</sub> (Sigma-Aldrich) and prepared by supported liquid extraction (SLE) using 96-well plates (Biotage, UK). Samples and calibrators were then subject to derivatization with DMEQ-TAD (Abcam, UK). LC–MS/MS analysis was conducted on a Shimadzu Nexera ultra-high-performance liquid chromatography (UPLC) system (Shimadzu Corporation, Kyoto, Japan) coupled to a Sciex QTrap 6500 quadrupole mass spectrometer (AB Sciex, Warrington, England). Liquid chromatography separation was carried out using a Raptor Fluorophenyl column (2.7  $\mu\text{m}$  100  $\text{\AA}$ , LC Column 100  $\times$  2.1 mm) (Thames Restek); the mobile phase was 2 mM ammonium formate in water with 0.1% formic acid (A) and 2 mM ammonium formate in methanol with 0.1% formic acid (B). For mass spectrometer analysis, ionization was performed by electrospray ionization (ESI) in positive ion mode and



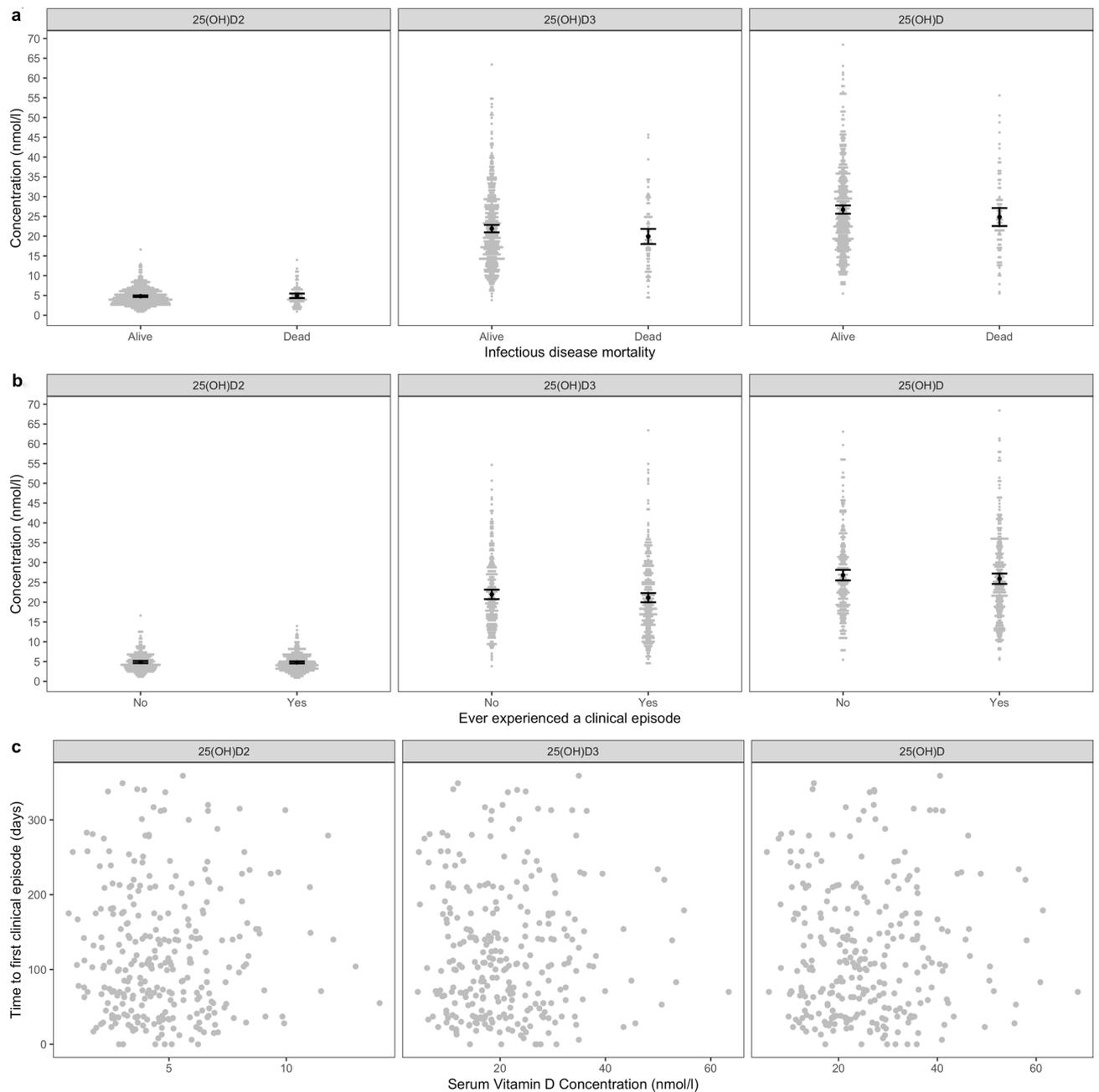
**Figure 6.** Estimates of effect sizes for the environmental level factors affecting serum 25(OH)D concentration after accounting for the calf level factors. Average model estimates are calculated for variables in the top model set (models with a cumulative Akaike weight  $\leq 0.95$ , Supplementary Table 2 and Supplementary Fig. 3). Error bars represent the 95% confidence intervals. Factors are considered important if their confidence intervals do not span one, as indicated by the dotted vertical line.



**Figure 7.** Estimates of effect sizes for the effect of wet/dry season on serum 25(OH)D concentration after accounting for the calf level factors. Error bars represent the 95% confidence intervals. Average model estimates are calculated for variables in the top model set (models with a cumulative Akaike weight  $\leq 0.95$ , Supplementary Table 3 and Supplementary Fig. 4). Factors are considered important if their confidence intervals do not span one, as indicated by the dotted vertical line.

multiple reaction monitoring (MRM) was used to monitor and quantify derivatized standards and endogenous vitamin D analytes, 25(OH)D<sub>2</sub> and 25(OH)D<sub>3</sub>, and internal standards, d<sub>3</sub>-25(OH)D<sub>2</sub> and <sup>13</sup>C<sub>5</sub>-25(OH)D<sub>3</sub>. Total 25(OH)D is defined as the sum of 25(OH)D<sub>2</sub> and 25(OH)D<sub>3</sub>.

**Genotyping of the cattle and SNP quality control.** All the cattle were genotyped with the 50 K Illumina BovineSNP50 beadchip v.1. The beadchip contains 55,777 SNPs spread evenly throughout the genome before quality control. Quality control was applied to all SNP data prior to analysis using PLINK 2.0<sup>69,70</sup>. SNPs with a minor allele frequency of  $< 0.01$  or a call rate of  $< 90\%$  were removed. In addition, individuals with a call rate of  $< 90\%$  were removed and only autosomal SNPs were included in the analysis. This left 40,405 autosomal variants and 518 cattle which passed quality control checks.

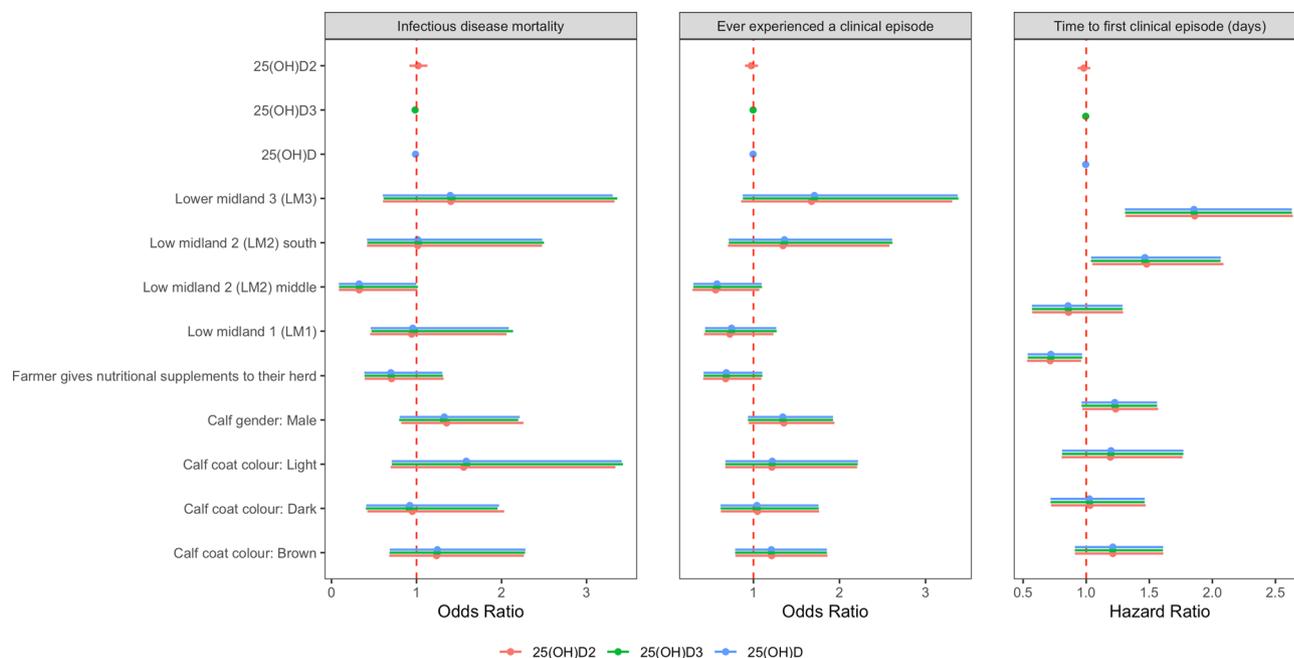


**Figure 8.** Serum 25(OH)D concentration and infectious disease mortality and morbidity by one year of age. Black points represent the mean with 95% confidence interval. **(a)** Infectious disease mortality; **(b)** experiencing a clinical episode by 51 weeks of age and **(c)** time to first clinical episode vs. serum 25(OH)D concentration.

**Statistical analysis.** The distributions of the total 25(OH)D, 25(OH)D<sub>2</sub> and 25(OH)D<sub>3</sub> were examined and the difference in median concentrations of 25(OH)D<sub>2</sub> and 25(OH)D<sub>3</sub> within samples were tested with a paired Wilcoxon signed rank test, as well as their correlation (Spearman rank correlation). Unless otherwise stated, all analysis was carried out using R version 3.6.0<sup>71</sup> and the figures were created using the R package *ggplot2*<sup>72</sup>.

**Association with calf level factors.** The calf level factors associated with serum 25(OH)D concentration was investigated using multi-model inference with model averaging. Calf level factors which were initially thought to be associated with serum 25(OH)D concentration were: calf gender; dam and calf coat colour, coat pattern and hair length; calf weight at recruitment; calf rectal temperature; dam body condition score and whether or not the farmer gives nutritional supplements to their herd.

Correlation between calf level variables were assessed using Pearson's product moment correlation for the continuous variables and chi-squared tests for the categorical variables. Since calf coat colour was correlated with variables which described both the dam's and the calf's coats, we avoided collinearity issues amongst the explanatory variables by only including calf coat colour in the global model to represent the coat descriptions.



**Figure 9.** Estimates of effect sizes for the association between infectious disease mortality and morbidity by one year of age and serum 25(OH)D concentration after accounting for calf level factors and vitamin D. Factors are considered important if their confidence intervals do not span one, as indicated by the dotted vertical line.

	Heritability	$V_A$	$V_R$
25(OH)D2	62.6% (35.3%)	0.20 (0.17)	0.12 (0.08)
25(OH)D3	70.1% (30.2%)	0.23 (0.16)	0.10 (0.07)
25(OH)D	71.1% (30.0%)	0.20 (0.13)	0.08 (0.06)

**Table 2.** Heritability and variance components of serum 25(OH)D concentration. The brackets represent the standard errors on the estimates.  $V_A$  additive genetic variance;  $V_R$  residual variation.

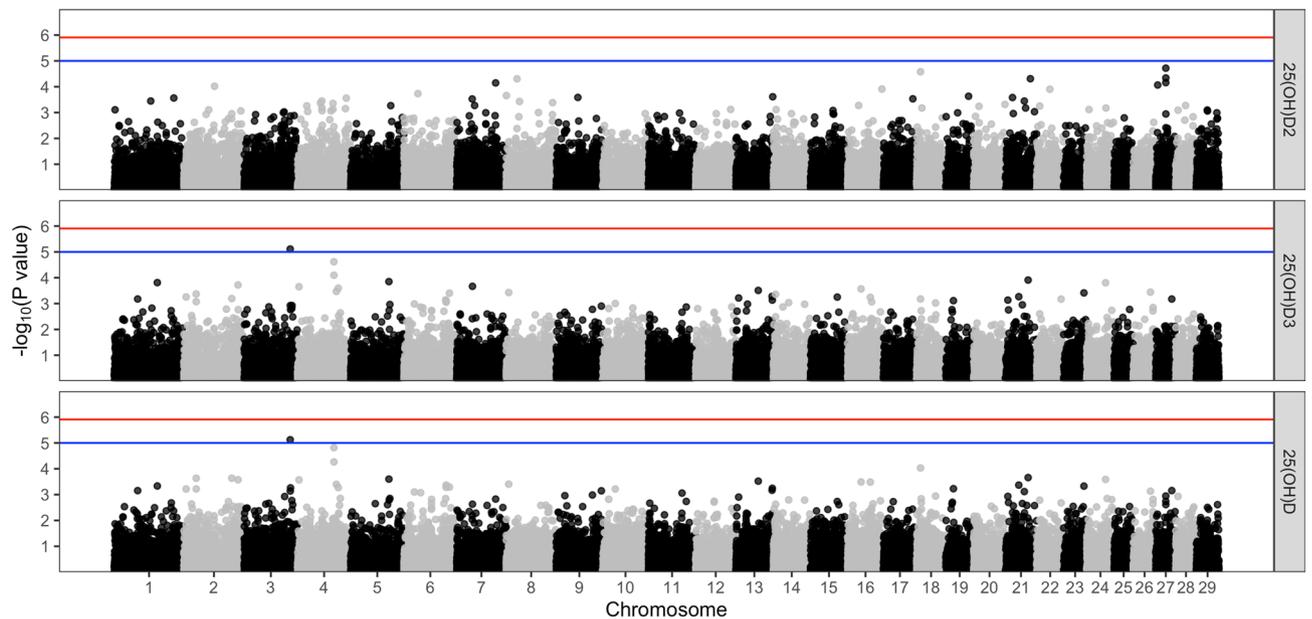
No other variables were correlated. Variables which were associated with any of the serum vitamin D metabolite concentrations with a P value < 0.2 was then included in the global model.

The global model for each metabolite was a generalised linear model fitted by restricted maximum likelihood (REML), with a gamma error distribution and log link function. The variables which passed the above selection criteria and were included as fixed effects in all the global model were calf gender, calf coat colour and whether or not the farmer gives nutritional supplements to their herd. Agroecological zone was included in all models as a fixed effect to account for similarity between cattle coming from the same AEZ. The global model was checked for variance inflation factors using the *vif* function of the *car* package<sup>73</sup>; no variables had a generalised variance inflation factor ( $GVI^{1/(2 \cdot df)}$ ) greater than the threshold value of 2<sup>74</sup>.

Model averaging was carried out using the ‘MuMIn’ package in R version 3.6.0<sup>71,75</sup> on the model set generated from the global model selecting all models were the cumulative Akaike weight less than or equal to 0.95. Parameter estimates and their confidence intervals was generated from this model set. The model averaging results were similar to those results produced from models without averaging.

**Association with environmental level factors.** Historic weather records for Busia, Kenya was obtained from <https://www.worldweatheronline.com>. Data was only available from June 2008 onwards, however the recruitment visits to the calves began in October 2007, therefore a large proportion of cattle are missing environmental data. In total we have complete records for 358 (68%) of individuals. The historic weather data consisted of number of hours of sunshine per day and percentage of cloud cover. This information was collapsed into monthly averages for the month prior to the visit at which the serum vitamin D sample was tested, to account for the weather variation the dam experienced before having the calf.

The calf level model described above, was extend to include the environmental level data. Therefore, the fixed effects included in the environmental model were: calf gender, calf coat colour and whether or not the farmer gives nutritional supplements to their herd; average total monthly sun hours prior to visit and average percentage cloud cover in month prior to visit (%). Agroecological zone was included in all models as a fixed effect to account



**Figure 10.** Manhattan plot of the genome wide association between SNPs and serum 25(OH)D concentration after accounting for Calf sex, calf coat colour, nutritional supplements, AEZ and the first 3 principle components as covariates. The blue line represents the suggestive significance line of  $< 10^{-5}$  and the red line represents the genome wide significance threshold, using a Bonferroni correction of  $< 10^{-6}$ .

for similarity between cattle coming from the same AEZ. The same methods of model selection, validation and averaging was used as the calf level model.

Broadly, the dry season in Kenya occurs between January to March and warm, dry period also occurs between July to October, whilst the wetter months are between April to June and November to December. The season in which the sample was taken was defined and the same modelling procedure as described above was used to evaluate the effect of the wet/dry season on serum 25(OH)D concentration. Since this broad definition of the wet/dry season does not rely on historic weather records, season could be defined for all 518 cattle.

**Association with infectious disease mortality and morbidity.** To investigate the relationship between infectious disease mortality by 1 year of age and 25(OH)D concentration a generalised linear model with a binomial error distribution and logit link functions was used. Serum 25(OH)D concentration was included as a continuous fixed effect in the model along with the calf level variables (calf coat colour, calf sex, nutritional supplements) and AEZ. A separate model was built for each vitamin D measurement.

Infectious disease morbidity was measured using two methods. Firstly, it was measured as whether or not the IDEAL veterinary clinicians recorded the calf to have ever experienced a clinical episode during its enrolment in the IDEAL study. This was modelled in the same way as infectious disease mortality, with a generalised linear model with a binomial error distribution and logit link functions was used. Serum 25(OH)D concentration was fitted as a continuous fixed effect.

Secondly, infectious disease morbidity was measured as the time to when the IDEAL veterinary clinicians first recorded that a calf experienced a clinical episode. In the event that cattle died from infectious causes before experiencing a clinical episode, the date of death was counted as the time to first clinical episode as it is an extreme case of a clinical illness. The time to first clinical episode was modelled using a Cox proportional hazards regression model. Like the other infectious disease morbidity model, 25(OH)D concentration was included as a continuous fixed effect in the model along with the calf level variables (calf coat colour, calf sex, nutritional supplements) and AEZ. A separate model was built for each vitamin D measurement. The hazard ratio was used to determine the effect of one unit change in serum 25(OH)D concentration on the risk of experiencing a clinical episode.

The association with infectious disease mortality and morbidity analysis was repeated with vitamin D included as a categorical variable based on tertiles in case the relationship between serum 25(OH)D concentration and health outcomes was not linear.

**Heritability of vitamin D.** To calculate the heritability of the serum 25(OH)D concentration, we first removed cattle which were related with more than a second-degree relative (corresponding to a KING kinship coefficient of 0.25) using PLINK<sup>69,70</sup>. This left 494 individuals for analysis. The variation in 25(OH)D concentration was partitioned into the additive genetic and the residual variance components using generalised linear mixed models with gamma error distributions and log link functions in ASReml version 3<sup>76</sup>. Calf sex, calf coat colour, nutritional supplements and AEZ were included as fixed effects in the model to account for the potential confounding effects of these factors.

The variance of the additive genetic effect,  $V_A$  (the genetic cause of variation in 25(OH)D concentration) was estimated by including individual identity linked with the genomic-relationship matrix (GRM) as a random effect in the model. The GRM was created constructed from the inverse of the IBS matrix which was created using the R package *SNPRelate*<sup>77</sup>. The remaining variation, is the residual error. It was assumed that the random effect, followed a normal distribution,  $a \sim N(0, \sigma_a^2 G)$ , where  $\sigma_a^2$  is the additive genetic variance and G is the genomic relation matrix described above; the residual error  $e \sim N(0, \sigma_e^2 I)$  with residual variance  $\sigma_e^2$  and identity matrix, I.

The narrow-sense heritability of serum 25(OH)D concentration ( $h^2$ ) was defined as the proportion of phenotypic variance explained by the additive genetic variance,  $h^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_e^2)$ . It describes the extent to which differences between individuals are determined by additive genetic effects<sup>78</sup>. The heritability analyses were carried out in ASReml version 3.0 as previously described<sup>76</sup>. Note, these estimates are likely to be inflated as they do not include any population structure that may be confounded with genetic effects. It was not possible to include these in the model due to the small numbers of individuals genotyped.

**Genome wide association with serum 25(OH)D concentration.** A genome wide association study (GWAS) was carried out on the quality controlled dataset using the GMMAT package in R<sup>71,79</sup>. Generalised linear models fitted with a gamma error distribution and log link functions was used to evaluate the association between serum 25(OH)D concentration and each SNP. Calf sex, calf coat colour, nutritional supplements, AEZ and the first three principle components where included as covariates in the model to account for population structure and confounding factors. A Bonferroni correction was applied to the genome-wide significance threshold, calculated as  $-\log_{10}(p) = -\log_{10}(0.05/40,405) = 5.91$ .  $p = 0.05/40,405 = 1^{-6}$ . SNPs which were associated with vitamin D with a p-value above this threshold where considered to be of interest.

**Ethics declarations.** The IDEAL project was reviewed and approved by the University of Edinburgh Ethics Committee (reference number OS 03-06) and also by the Institute Animal Care and Use Committee of the International Livestock Research Institute, Nairobi. Standard techniques were used to collect blood and faecal samples for diagnosis and identification of disease and infecting pathogens. The calves were restrained by professional animal health assistants or by qualified veterinary surgeons. A veterinary surgeon was available to examine any calf falling sick during the course of the study. Any calf that was in severe distress due to trauma or disease was humanely euthanized by intravenous injection of sodium pentobarbital by a veterinary surgeon. All participating farmers gave informed consent in their native language before recruiting their animals into the study. All methods were carried out in accordance with relevant guidelines and regulations.

Received: 29 May 2020; Accepted: 19 October 2020

Published online: 25 November 2020

## References

- Elder, C. J. & Bishop, N. J. Rickets. *Lancet* **383**(9929), 1665–1676 (2014).
- Dittmer, K. E. & Thompson, K. G. Vitamin D metabolism and rickets in domestic animals: A review. *Vet. Pathol.* **48**(2), 389–407 (2011).
- Hymoller, L. & Jensen, S. K. Vitamin D(3) synthesis in the entire skin surface of dairy cows despite hair coverage. *J. Dairy Sci.* **93**(5), 2025–2029 (2010).
- Mellanby E. Nutrition classics. *Lancet* **1**, 407–12 (1919) (an experimental investigation of rickets. Edward Mellanby. *Nutr. Rev.* **34**(11), 338–40, 1976).
- Provvedini, D. M., Tsoukas, C. D., Defetos, L. J. & Manolagas, S. C. 1,25-dihydroxyvitamin D3 receptors in human leukocytes. *Science* **221**(4616), 1181–1183 (1983).
- Holick, M. F. Vitamin D deficiency. *N. Engl. J. Med.* **357**(3), 266–281 (2007).
- Baeke, F., Takiishi, T., Korf, H., Gysemans, C. & Mathieu, C. Vitamin D: Modulator of the immune system. *Curr. Opin. Pharmacol.* **10**(4), 482–496 (2010).
- Besusso, D. *et al.* 1,25-Dihydroxyvitamin D-conditioned CD11c+ dendritic cells are effective initiators of CNS autoimmune disease. *Front. Immunol.* **6**, 575 (2015).
- Saul, L. *et al.* 1,25-Dihydroxyvitamin D3 restrains CD4(+) T cell priming ability of CD11c(+) dendritic cells by upregulating expression of CD31. *Front. Immunol.* **10**, 600 (2019).
- Afzal, S., Brondum-Jacobsen, P., Bojesen, S. E. & Nordestgaard, B. G. Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts. *BMJ* **349**, g6330 (2014).
- Chowdhury, R. *et al.* Vitamin D and risk of cause specific death: Systematic review and meta-analysis of observational cohort and randomised intervention studies. *BMJ* **348**, g1903 (2014).
- Schottker, B. *et al.* Vitamin D and mortality: Meta-analysis of individual participant data from a large consortium of cohort studies from Europe and the United States. *BMJ* **348**, g3656 (2014).
- Ganmaa, D., Munkhzul B, Fawzi W, Spiegelman D, Willett WC, Bayasgalan P, *et al.* High-dose vitamin D3 during tuberculosis treatment in Mongolia. A randomized controlled trial. *Am. J. Respir. Crit. Care Med.* **196**(5), 628–637 (2017).
- Aibana, O. *et al.* Vitamin D status and risk of incident tuberculosis disease: A nested case-control study, systematic review, and individual-participant data meta-analysis. *PLoS Med.* **16**(9), e1002907 (2019).
- Wu, H. X. *et al.* Effects of vitamin D supplementation on the outcomes of patients with pulmonary tuberculosis: A systematic review and meta-analysis. *BMC Pulm. Med.* **18**(1), 108 (2018).
- Jolliffe, D. A. *et al.* Vitamin D supplementation to prevent asthma exacerbations: A systematic review and meta-analysis of individual participant data. *Lancet Respir. Med.* **5**(11), 881–890 (2017).
- Munger, K. L., Levin, L. I., Hollis, B. W., Howard, N. S. & Ascherio, A. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. *JAMA* **296**(23), 2832–2838 (2006).
- Rhead, B. *et al.* Mendelian randomization shows a causal effect of low vitamin D on multiple sclerosis risk. *Neurol. Genet.* **2**(5), e97 (2016).

19. Manson, J. E. *et al.* Vitamin D supplements and prevention of cancer and cardiovascular disease. *N. Engl. J. Med.* **380**(1), 33–44 (2019).
20. Pittas, A. G. *et al.* Vitamin D supplementation and prevention of type 2 diabetes. *N. Engl. J. Med.* **381**(6), 520–530 (2019).
21. Titmarsh, H. *et al.* Vitamin D status predicts 30 day mortality in hospitalised cats. *PLoS ONE* **10**(5), e0125997 (2015).
22. Jaffey, J. A., Backus, R. C., McDaniel, K. M. & DeClue, A. E. Serum vitamin D concentrations in hospitalized critically ill dogs. *PLoS ONE* **13**(3), e0194062 (2018).
23. Titmarsh, H., Gow, A.G., Kilpatrick, S., Sinclair, J., Hill, T., Milne, E., *et al.* Association of vitamin D status and clinical outcome in dogs with a chronic enteropathy. *J. Vet. Intern. Med.* (2015).
24. Allenspach, K., Rizzo, J., Jergens, A. E. & Chang, Y. M. Hypovitaminosis D is associated with negative outcome in dogs with protein losing enteropathy: A retrospective study of 43 cases. *BMC Vet. Res.* **13**(1), 96 (2017).
25. Titmarsh, H. F. *et al.* Low vitamin D status is associated with systemic and gastrointestinal inflammation in dogs with a chronic enteropathy. *PLoS ONE* **10**(9), e0137377 (2015).
26. Titmarsh, H.F., Cartwright, J.A., Kilpatrick, S., Gaylor, D., Milne, E.M., Berry, J.L., *et al.* Relationship between vitamin D status and leukocytes in hospitalised cats. *J. Feline Med. Surg.* (2016).
27. Handel, I. *et al.* Vitamin D status predicts reproductive fitness in a wild sheep population. *Sci. Rep.* **6**, 18986 (2016).
28. Zhou, P. *et al.* Investigation of relationship between vitamin D status and reproductive fitness in Scottish hill sheep. *Sci. Rep.* **9**(1), 1162 (2019).
29. Corripio-Miyar, Y., Mellanby, R. J., Morrison, K. & McNeilly, T. N. 1,25-Dihydroxyvitamin D3 modulates the phenotype and function of Monocyte derived dendritic cells in cattle. *BMC Vet. Res.* **13**(1), 390 (2017).
30. Yue, Y., Hymoller, L., Jensen, S. K., Lauridsen, C. & Purup, S. Effects of vitamin D and its metabolites on cell viability and *Staphylococcus aureus* invasion into bovine mammary epithelial cells. *Vet. Microbiol.* **203**, 245–251 (2017).
31. Garcia-Barragan, A., Gutierrez-Pabello, J. A. & Alfonso-Silva, E. Calcitriol increases nitric oxide production and modulates microbicidal capacity against *Mycobacterium bovis* in bovine macrophages. *Comp. Immunol. Microbiol. Infect. Dis.* **59**, 17–23 (2018).
32. Waters, W. R. *et al.* Modulation of *Mycobacterium bovis*-specific responses of bovine peripheral blood mononuclear cells by 1,25-dihydroxyvitamin D(3). *Clin. Diagn. Lab. Immunol.* **8**(6), 1204–1212 (2001).
33. Lippolis, J. D., Reinhardt, T. A., Sacco, R. A., Nonnecke, B. J. & Nelson, C. D. Treatment of an intramammary bacterial infection with 25-hydroxyvitamin D(3). *PLoS ONE* **6**(10), e25479 (2011).
34. Merriman, K. E., Poindexter, M. B., Kweh, M. E., Santos, J. E. P. & Nelson, C. D. Intramammary 1,25-dihydroxyvitamin D3 treatment increases expression of host-defense genes in mammary immune cells of lactating dairy cattle. *J. Steroid Biochem. Mol. Biol.* **173**, 33–41 (2017).
35. Stabel, J. R., Reinhardt, T. A. & Hempel, R. J. Short communication: Vitamin D status and responses in dairy cows naturally infected with *Mycobacterium avium* ssp. paratuberculosis. *J. Dairy Sci.* **102**(2), 1594–1600 (2019).
36. de Clare Bronsvoort, B. M. *et al.* Design and descriptive epidemiology of the Infectious Diseases of East African Livestock (IDEAL) project, a longitudinal calf cohort study in western Kenya. *BMC Vet. Res.* **9**, 171 (2013).
37. Hunt, S.E., McLaren, W., Gil, L., Thormann, A., Schuilenburg, H., Sheppard, D., *et al.* Ensembl variation resources. *Database* (2018).
38. Jiang, X. *et al.* Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat. Commun.* **9**(1), 260 (2018).
39. Manousaki, D., Mitchell, R., Dudding, T., Haworth, S., Harroud, A., Forgetta, V., *et al.* Genome-wide association study for vitamin D levels reveals 69 independent loci. *Am. J. Hum. Genet.* (2020).
40. Titmarsh, H. *et al.* Association of vitamin D status and clinical outcome in dogs with a chronic enteropathy. *J. Vet. Intern. Med.* **29**(6), 1473–1478 (2015).
41. Aspelund, T. *et al.* Effect of genetically low 25-hydroxyvitamin D on mortality risk: Mendelian randomization analysis in 3 large European cohorts. *Nutrients.* **11**(1), 1 (2019).
42. Huang, T. *et al.* Vitamin D and cause-specific vascular disease and mortality: a Mendelian randomisation study involving 99,012 Chinese and 106,911 European adults. *BMC Med.* **17**(1), 160 (2019).
43. Ong, J. S. *et al.* Association of vitamin D levels and risk of ovarian cancer: A Mendelian randomization study. *Int. J. Epidemiol.* **45**(5), 1619–1630 (2016).
44. Reid, D. *et al.* The relation between acute changes in the systemic inflammatory response and plasma 25-hydroxyvitamin D concentrations after elective knee arthroplasty. *Am. J. Clin. Nutr.* **93**(5), 1006–1011 (2011).
45. Waldron, J. L. *et al.* Vitamin D: a negative acute phase reactant. *J Clin Pathol.* **66**(7), 620–622 (2013).
46. Nonnecke, B. J. *et al.* Acute phase response elicited by experimental bovine diarrhea virus (BVDV) infection is associated with decreased vitamin D and E status of vitamin-replete preruminant calves. *J. Dairy Sci.* **97**(9), 5566–5579 (2014).
47. Nelson, C. D. *et al.* Assessment of serum 25-hydroxyvitamin D concentrations of beef cows and calves across seasons and geographical locations. *J. Anim. Sci.* **94**(9), 3958–3965 (2016).
48. Saliba, W., Barnett, O., Stein, N., Kershenbaum, A. & Rennert, G. The longitudinal variability of serum 25(OH)D levels. *Eur. J. Intern. Med.* **23**(4), e106–e111 (2012).
49. Major, J. M. *et al.* Variability and reproducibility of circulating vitamin D in a nationwide U.S. population. *J. Clin. Endocrinol. Metab.* **98**(1), 97–104 (2013).
50. Jorde, R. *et al.* Tracking of serum 25-hydroxyvitamin D levels during 14 years in a population-based study and during 12 months in an intervention study. *Am. J. Epidemiol.* **171**(8), 903–908 (2010).
51. McKibben, R. A. *et al.* Factors associated with change in 25-hydroxyvitamin D levels over longitudinal followup in the ARIC study. *J. Clin. Endocrinol. Metab.* **1**, jc20151711 (2015).
52. van Schoor, N. M. *et al.* Longitudinal changes and seasonal variations in serum 25-hydroxyvitamin D levels in different age groups: Results of the Longitudinal Aging Study Amsterdam. *Osteoporos. Int.* **25**(5), 1483–1491 (2014).
53. Liu, X. *et al.* Longitudinal trajectory of vitamin D status from birth to early childhood in the development of food sensitization. *Pediatr. Res.* **74**(3), 321–326 (2013).
54. Baiz, N. *et al.* Cord serum 25-hydroxyvitamin D and risk of early childhood transient wheezing and atopic dermatitis. *J. Allergy Clin. Immunol.* **133**(1), 147–153 (2014).
55. Lai, S. H. *et al.* Low cord-serum 25-hydroxyvitamin D levels are associated with poor lung function performance and increased respiratory infection in infancy. *PLoS ONE* **12**(3), e0173268 (2017).
56. Sauder, K. A. *et al.* Cord blood vitamin D levels and early childhood blood pressure: The healthy start study. *J. Am. Heart Assoc.* **8**(9), e011485 (2019).
57. Cetinkaya, M. *et al.* Lower vitamin D levels are associated with increased risk of early-onset neonatal sepsis in term infants. *J. Perinatol.* **35**(1), 39–45 (2015).
58. Behera, C.K., Sahoo, J.P., Patra, S.D., Jena, P.K. Is lower vitamin D level associated with increased risk of neonatal sepsis? A prospective cohort study. *Indian J. Pediatr.* (2020).
59. Sacco, R. E. *et al.* Differential expression of cytokines in response to respiratory syncytial virus infection of calves with high or low circulating 25-hydroxyvitamin D3. *PLoS ONE* **7**(3), e33074 (2012).
60. Yue, Y., Hymoller, L., Jensen, S. K. & Lauridsen, C. Effect of vitamin D treatments on plasma metabolism and immune parameters of healthy dairy cows. *Arch. Anim. Nutr.* **72**(3), 205–220 (2018).

61. Hidirolou, M., Williams, C. J. & Proulx, J. G. Plasma vitamin D3 response in cattle and sheep exposed to ultraviolet radiation. *Int. J. Vitam. Nutr. Res.* **55**(1), 41–46 (1985).
62. Casas, E., Lippolis, J. D., Kuehn, L. A. & Reinhardt, T. A. Seasonal variation in vitamin D status of beef cattle reared in the central United States. *Domest. Anim. Endocrinol.* **52**, 71–74 (2015).
63. Hymoller, L., Jensen, S. K., Kaas, P. & Jakobsen, J. Physiological limit of the daily endogenous cholecalciferol synthesis from UV light in cattle. *J. Anim. Physiol. Anim. Nutr. (Berl)*. **101**(2), 215–221 (2017).
64. Jolliffe, D. A. *et al.* Vitamin D to prevent exacerbations of COPD: Systematic review and meta-analysis of individual participant data from randomised controlled trials. *Thorax* **74**(4), 337–345 (2019).
65. Martineau, A. R. *et al.* Vitamin D supplementation to prevent acute respiratory tract infections: Systematic review and meta-analysis of individual participant data. *BMJ* **356**, i6583 (2017).
66. Nicholson, M., Butterworth, M.H. A guide to condition scoring of zebu cattle: ILRI (aka ILCA and ILRAD) (1986).
67. Callaby, R., Pendarovski, C., Thumbi, S.M., Van Wyk, I., Mbole-Kariuki, M.N., Kiara, H., *et al.* The infectious diseases of East African livestock (IDEAL) project database. *Nat. Sci. Data* (**submitted**).
68. Hurst, E.A., Homer, N.Z., Gow, A.G., Clements, D.N., Evans, H., Gaylor, D., *et al.* Vitamin D status is seasonally stable in northern European dogs. *Vet. Clin. Pathol.* (2020).
69. Purcell, S., Chang, C. PLINK 2.0.
70. Chang, C. C. *et al.* Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*. **4**(1), 1 (2015).
71. R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, Vienna, 2019).
72. Wickham H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, New York, 2016).
73. Fox, J. & Weisberg, S. *An R companion to applied regression* (Sage, Thousand Oaks, 2019).
74. Zuur, A. F., Ieno, E. N. & Elphick, C. S. A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* **1**(1), 3–14 (2010).
75. Bartoń K. MuMIn: Multi-Model Inference. *R package version 1.43.15* (2019).
76. Gilmour, A., Gogel, B., Cullis, B., Thompson, R., Butler, D., Cherry, M., *et al.* *ASReml user guide release 3.0*. (VSN Int Ltd, 2008).
77. Zheng, X. *et al.* A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**(24), 3326–3328 (2012).
78. Hill, W. G. & Mackay, T. F. DS Falconer and introduction to quantitative genetics. *Genetics* **167**(4), 1529–1536 (2004).
79. Chen H. *GMMAT: Generalized linear Mixed Model Association Tests Version 1.1. 2.* (2019).
80. Turner S. *qqman: Q-Q and Manhattan Plots for GWAS Data.* (2017).

## Acknowledgements

This work would not be possible without samples collected by the Infectious Diseases of East African Livestock (IDEAL) project, a collaborative project between the University of Edinburgh, University of Pretoria, University of Nottingham and the International Livestock Research Institute (ILRI), Nairobi, Kenya. We would like to thank the IDEAL consortium for providing the samples. We gratefully acknowledge the work of the IDEAL project field and laboratory teams, assistance from the Kenya Department of Veterinary Services in the running of the IDEAL project. In addition, we would like to thank Maurine C. Makau who retrieved the serum samples from the freezers, aliquotted them, packed them and arranged shipment and Elizabeth A. J. Cook who arranged the export permits.

## Authors contributions

Conceived and designed experiments: R.J.M and B.M.B. Performed the experiments: E.H., P.T., I.H. Analysed the data: R.C. Wrote and reviewed the paper: R.J.M. and R.C. All authors reviewed the final paper.

## Funding

This research was funded in part by the Bill & Melinda Gates Foundation and with UK aid from the UK Government's Department for International Development (Grant Agreement OPP1127286) under the auspices of the Centre for Tropical Livestock Genetics and Health (CTLGH), established jointly by the University of Edinburgh, SRUC (Scotland's Rural College), and the International Livestock Research Institute. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation nor the UK Government. In addition, the IDEAL project was funded by a Wellcome Trust Programme Grant (WT079445). Barend MDeC Bronsvort was supported through BBSRC through the Institute Strategic Programme funding (BB/J004235/1 and BB/P013740/1).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-77209-5>.

**Correspondence** and requests for materials should be addressed to R.J.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020